



THE
ONTARIO WATER RESOURCES
COMMISSION

REVIEW

of

WATER QUALITY STATION

DENSITY STATISTICAL PROCEDURES

1968

MOE
REV
ANVE

c.1
a aa

Copyright Provisions and Restrictions on Copying:

This Ontario Ministry of the Environment work is protected by Crown copyright (unless otherwise indicated), which is held by the Queen's Printer for Ontario. It may be reproduced for non-commercial purposes if credit is given and Crown copyright is acknowledged.

It may not be reproduced, in all or in part, for any commercial purpose except under a licence from the Queen's Printer for Ontario.

For information on reproducing Government of Ontario works, please contact ServiceOntario Publications at copyright@ontario.ca

REVIEW OF WATER QUALITY STATION

DENSITY STATISTICAL PROCEDURES

1968

M. D. Palmer and G. K. Sato,
Water Quality Surveys Branch,
Division of Sanitary Engineering.

WQR-2

REVIEW OF WATER QUALITY STATION

DENSITY STATISTICAL PROCEDURES

ABSTRACT

Data on selected water quality parameters measured during 1966 and 1967 at a small number of judiciously chosen stations in Lake Ontario were statistically compared to test whether each station was producing unique water quality information. The stations were compared utilizing established parametric and non-parametric methods both on a pairs and multiple station basis. As a result of this study, a larger project study is recommended involving the following:

1. analyzing an area of stations for station differences using total phosphates and pH employing the Mann-Whitney U test for pair testing and the Kruskal-Wallis one-way analysis of variance for multiple station testing.
2. installing a submersible recording type water quality meter to determine population distributions for dissolved oxygen, conductivity, pH and turbidity.

REVIEW OF WATER QUALITY STATION

DENSITY STATISTICAL PROCEDURES

TABLE OF CONTENTS

	PAGE NO.
INTRODUCTION	1
OUTLINE	2
PURPOSE	3
I SELECTION OF TEST PARAMETERS	3
II PAIR TESTING	9
Selection of Stations	11
Parametric Analysis	11
Normalizing Transforms	12
Non-Parametric Analysis.....	19
Discussion	25
Conclusions	27
III MULTIPLE STATION TESTING	28
Selection of Stations	28
Parametric Analysis of Variance	28
Non-Parametric Analysis of Variance	35
Discussion	36
Conclusions	37
SUMMARY OF CONCLUSIONS	38
APPENDIX I Computer Programs Developed	41
APPENDIX II Summary of Pairs Analysis	43
APPENDIX III Chloride Data	44

REVIEW OF WATER QUALITY STATION

DENSITY STATISTICAL PROCEDURES

INTRODUCTION

The Water Quality Surveys Branch has conducted a regular water quality monitoring survey along the Canadian shore of lakes Erie and Ontario since 1966. The purpose of these surveys is to obtain an inventory of water quality in the near shore areas and its relationship to material input sources. Over several years, the inventories can be employed to detect long term environmental changes as well as more detailed information on the cause and effect relationships. Obviously, the foundation of effective water management rests on a sound water quality inventory system. As such systems are expensive to maintain, it is important that the network of sampling stations employed is the most efficient one commensurate with existing survey techniques and information requirements. At the initiation of the survey program, a sampling grid was set-up based upon existing knowledge and intuition concerning the geographical location of land runoff, industrial and urban waste sources. Sampling locations thus established have been utilized during the past three years. Sufficient data has thus been collected to enable an analytical assessment of the existing grid to indicate what adjustments can be incorporated to make the grid more

efficient in describing the water quality variations by eliminating stations which are producing similar water quality information and introducing new stations where inadequate data is being accumulated.

A study was, therefore, initiated to test various known statistical techniques for their applicability to detect similarity and/or differences between the water quality information collected from different sampling stations. The testing of the statistical techniques was conducted on a limited basis by comparing pairs of stations or a small group of stations in a selected area to permit detailed investigation of each technique and its underlying assumptions. Further, it is known from other analytical treatments that the behaviour of various water quality parameters is not normally distributed probably due to non-random physical factors such as wind, currents, thermal stratification and variable loadings. Any statistical technique adopted must incorporate some allowances for these factors.

The problem of determining where more stations are required is not treated explicitly in this report although it is indirectly determined by establishing the variation in data between stations.

OUTLINE

The study is treated under three basic headings, namely selection of test parameters to be used; station pair-testing and multiple station testing.

Selection of parameters involves the identification of the most critical water quality parameters. This selection is based on the accuracy of determinations, the characteristics of the parameter in the environment and other studies.

Station pair-testing deals with the testing of various pairs of stations employing parametric, data transformation and non-parametric techniques with the appropriate discussions of various methods centering around the distributions of the parameters.

The multiple station testing deals primarily with parametric and non-parametric analysis of variance and the associated problems in their application.

PURPOSE

The objective of this work is to develop a reliable analytical technique for the evaluation of existing sampling grids and to eliminate stations which are not contributing significantly to the water quality information system.

I SELECTION OF TEST PARAMETERS

The task of selecting reliable water quality parameters which would indicate different water environmental characteristics is a difficult

one. Obviously, it must be a common parameter which has been sampled frequently with a reliable determination accuracy. Further, it must be a good indicator of pollution.

Accuracies in measuring selected water quality parameters during different survey years are presented in Table 1. It must be pointed out that the problem of accurately determining parameters in the lake environment using standard techniques is difficult due to the low concentrations encountered. Methods employed must be consistent and applicable on a production basis to be useful on surveys. Continuing work in the area of water quality analysis has, however, produced successively better techniques. Generally, there was an increased accuracy during 1968 surveys through improved instrumentation. Estimates of accuracies for 1966-67 determinations are approximate and in keeping with the techniques employed at the time. Improvement in the analytical accuracies between survey years introduces another complication into the analysis which may in part be overcome by judiciously selecting parameters. For instance, the accuracy of determining chlorides, alkalinity, dissolved oxygen (DO) and pH has remained constant throughout the survey years 1966-68. While the accuracies of determining ammonia, phenols and phosphates has improved since 1966. General improvements have also been made in phosphate determinations. The importance of the phosphate parameter for nutrient evaluation dictates its inclusion in any analytical system.

TABLE 1
WATER QUALITY ANALYSIS
ACCURACIES OF DETERMINATIONS
1966-7

Water Quality Parameter	Range	No. of Readings	Standard Deviations
Alkalinity ppm-CaCO ₃	93-100	52	5.09
Dissolved Oxygen (DO) ⁽¹⁾ percent sat.	94-130	135	2.3
Nitrogen components as N			
NH ₃ ⁽²⁾ ppm	0.03-0.67	-	No figures available large values too high low values too low
Kjeldahl ppm	-	-	± 5 to 20 percent error
NO ₃ and NO ₂ ppm	0.01-1.70	-	± 100 percent
pH Su	8.2-8.6	63	0.057
Phosphate as PO ₄			
Total ppm	0.02-0.40	-	± 5 to 20 percent
Phenols ⁽³⁾	-	-	Values below 5 not significant

(1) techniques for determining temperatures of deep samples were not reliable.

(2) 1966 results were very erratic and rejected.

(3) 1966 results were erratic, possibly due to contaminated samplers, and rejected.

TABLE 1
WATER QUALITY ANALYSIS
ACCURACIES OF DETERMINATIONS
1968

Water Quality Parameter	Range	No. of Readings	Standard Deviations
Alkalinity ppm-CaCO ₃	93-100	52	5.04
Dissolved Oxygen (DO) percent sat.	94-130	135	2.3
Nitrogen components as N			
NH ₃ ppm	0.03-0.67	51	0.0084
Kjeldahl ppm	0.1-1.6	54	0.058
NO ₃ and NO ₂ ppm	0.01-1.62	48	0.022
pH Su	8.2-8.6	63	0.057
Phosphate as PO ₄			
Total ppm	0.02-0.31	54	0.029
Soluble ppm	0.02-0.11	54	0.008
Phenols ppm	-	-	-

In recent studies conducted on Lake Ontario at Pickering and on Lake Erie at Nanticoke (Palmer, 1968) designed to determine the density of sampling stations required in areas close to shore, some enlightening facts concerning water quality parameters were discovered. Both surveys were conducted over a 24-hour period utilizing multiple sampling techniques. An analysis of variance of the results revealed that alkalinity, total phosphate and pH were the most sensitive indicators of differences between sampling points. This work also illustrated a significant diurnal variation in DO, pH, alkalinity and phosphate concentrations at all near shore stations. Greatest variations were noted between 5 p.m. and 3 a.m. Consequently, the time of sampling is important since diurnal variation in parameters will obviously affect the analysis.

The following parameters were selected for the development and testing of analysis techniques:

1. dissolved oxygen (DO)
2. phosphates
3. alkalinity

The rejection of pH as an indicator for this study was based upon the fact that pH produced conflicting results between two studies on sampling station density requirements (Palmer, 1968). Since the variation in chlorides were small, this parameter was also rejected (see Appendix IX).

Furthermore, it is appreciated that DO and phosphates are active components which are affected by bottom and/or atmospheric surface exchanges.

II PAIR TESTING

Analytical techniques for comparing two sampling stations for equality of means and variance (measure of the variation) are well established and documented. The techniques employed here closely follow the procedures outlined in:

- (1) Experimental Statistics, Handbook 91.
U. S. Department of Commerce. 1963
- (2) Engineering Statistics by A. H. Bowker
and G. J. Kleberman. Prentice-Hall 1959.
- (3) Non-Parametric Statistics by S. Siegel
MaGraw-Hill. 1956

Generally, analytical techniques for testing two stations are statistically strong, if adequate numbers are available. Consequently, it was decided to start by using pair testing appreciating that its application to a whole station system would be tedious and expensive in computer time. It is furthermore easier to identify the results for pairs and visualize what is happening. Analysis of variance techniques for multiple stations will be discussed later and in fact is really only an extension of the pair testing.

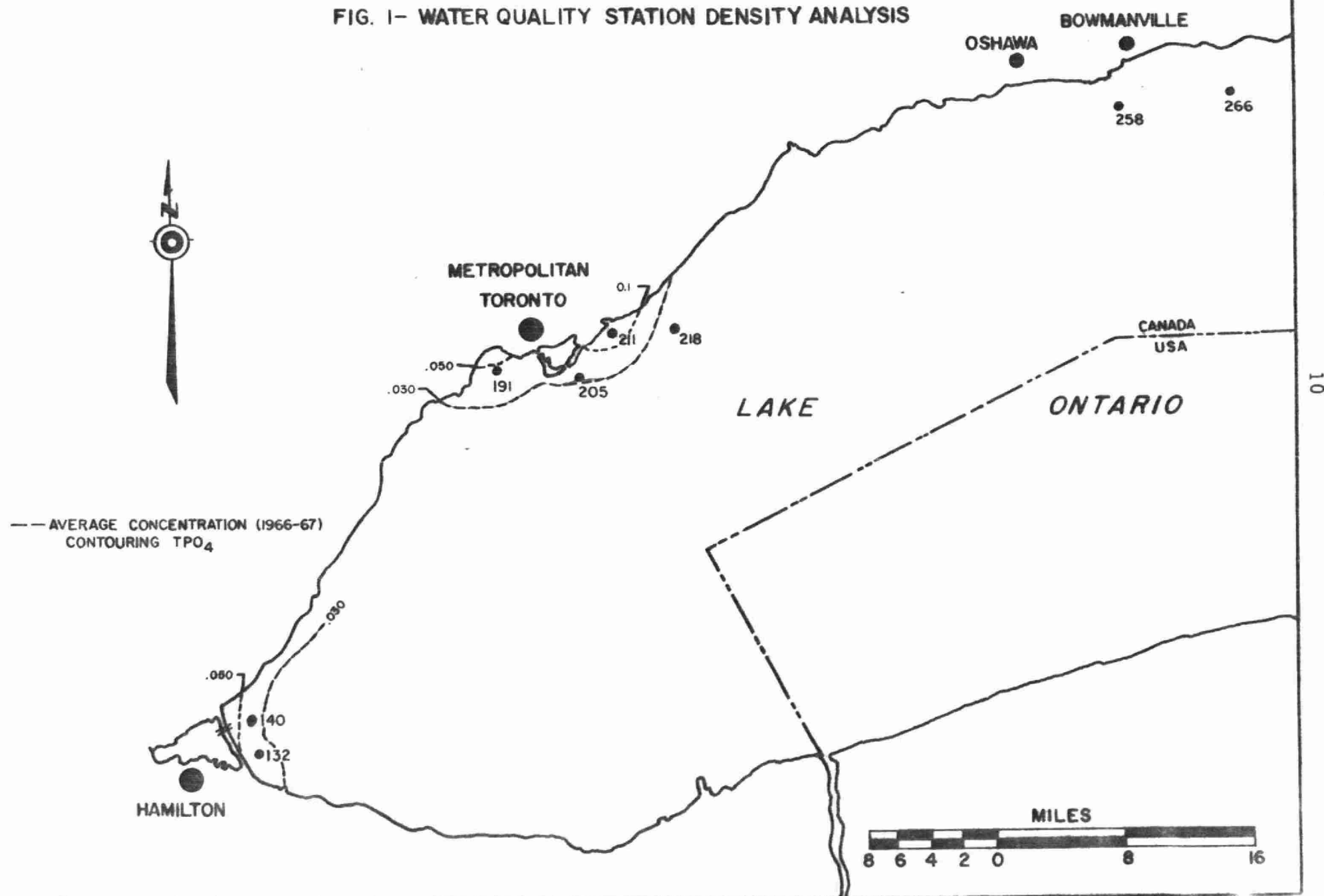
TABLE 2

PAIR TESTING

PARAMETRIC ANALYSIS, F AND T TESTS

Stations	Parameter	F (calc.)	γ_1/γ_2	F.025, α, β	Remarks	t.025, α, β	t (calc.)	t.025, γ	Remarks
132/140	DO	1.04	23/40	2.15	N.S.D.	2.003	0.2651		N.S.D.
	CaCO ₃	1.17	23/30	2.15	N.S.D.	2.003	-0.1448		N.S.D.
	S.PO ₄	2.36	28/31	2.33	S.D.		-48.1	2.021	S.D.
	T.PO ₄	2.81	29/22	2.28	S.D.		-33.1	2.021	S.D.
132/258	DO	1.65	23/30	2.15	N.S.D.	2.003	2.07		S.D.
	CaCO ₃	1.27	31/23	2.23	N.S.D.	2.003	-1.13		N.S.D.
	S.PO ₄	19.13	29/21	2.32	S.D.		-14.80	2.045	S.D.
	T.PO ₄	10.57	29/23	2.28	S.D.		8.79	2.045	S.D.

FIG. 1- WATER QUALITY STATION DENSITY ANALYSIS



Selection of Stations

There must be enough data available to make analysis feasible and statistically strong to permit the selection of representative stations for pair testing. Stations near the shore will obviously be affected by local sources or in some cases by a single source with variable loadings, while stations farther offshore in deeper water are, in most cases, affected by thermal stratification. Stations geographically separated by large distances are generally in different water movement areas. While it is impossible to account for all the variables mentioned, they can be minimized by utilizing the parameter concentration contouring produced by plotting long term average values. Such contours smooth out the effects of the variables at a station over the survey season. Station pairs selected for analysis were based on contour plots.

Parametric Analysis

The underlying assumption of this technique is that the parameters are normally distributed resorting to the Central Limit Theorem. Two pairs of stations numbers 132/140 and 132/258 (see Figure 1) were selected for analysis using the corresponding data for the years 1966 to 1967. The data was then tested for equality of variance (Bowker, 1959) employing an "F" test. Subsequently, the pairs were tested for equality of means employing a "t" test (Bowker, 1959) under two hypothesis: namely, with variances assumed equal and not equal (see Table 2).

From the concentration contouring, it was expected that Stations #132/140 would be similar and stations #132/258 would be different. Whereas both pairs were found to be significantly different in overall means and variances for phosphates and overall means for DO, this result was not reflected in the concentration contours. Consequently, the question arises concerning the underlying assumption of a normal distribution.

Normalizing Transforms

While it is difficult to determine the distribution of a variable without resorting to population studies, indications can sometimes be obtained by plotting large sample distributions, then resorting to normalizing transformations of the data.

Frequency analyses were performed on the data from the selected stations to determine an estimate of the distribution of parameters. Most of the distributions (see Figures 2, 3, 4 and 5) were found to be non-normal (NN). Scales on which some of these parameters were measured do not appear suitable for parametric statistics due to the assumed condition of normality of the population. To alleviate this situation, a transformation (change of scale) can be applied to the raw data so that parametric tests could be validly performed.

To explicitly determine an applicable transformation, it is necessary to measure a population distribution or at least a very large

FIGURE 2
TOTAL PHOSPHATES
STATION # 258
1966-67

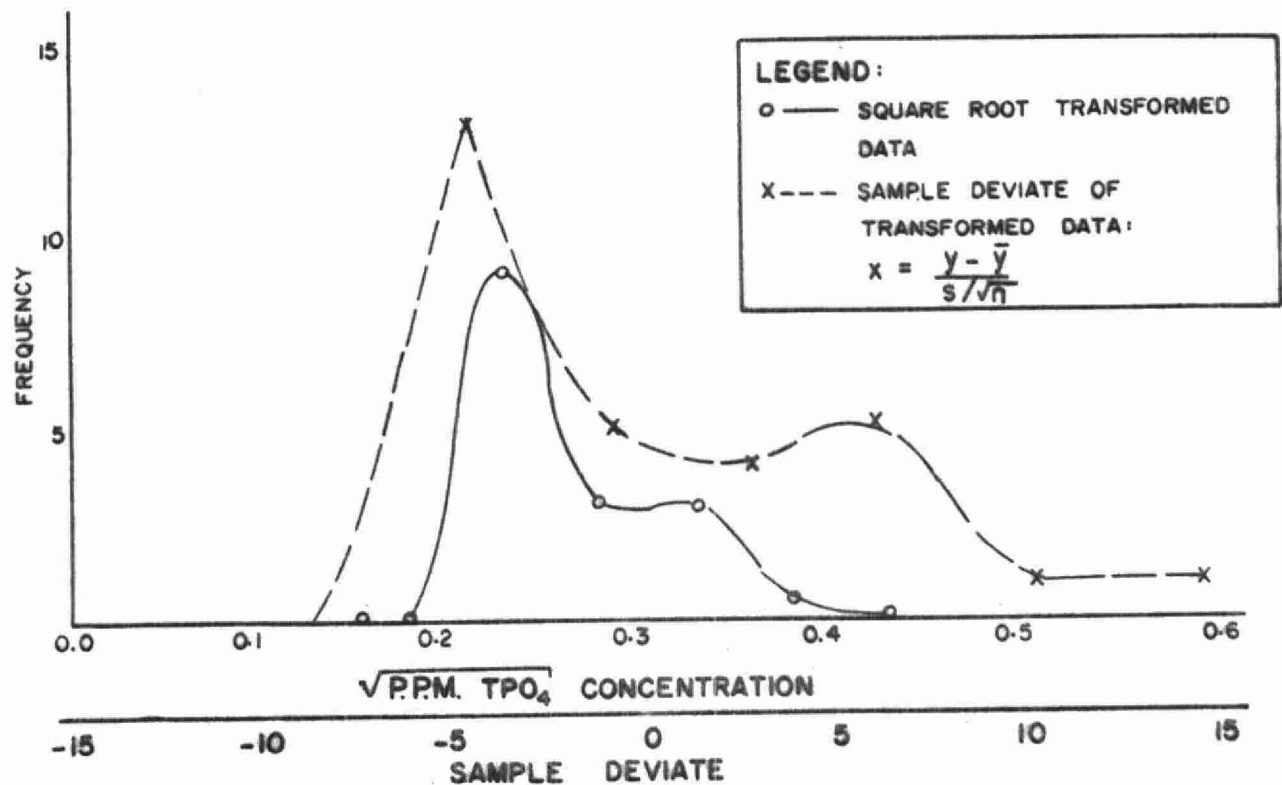
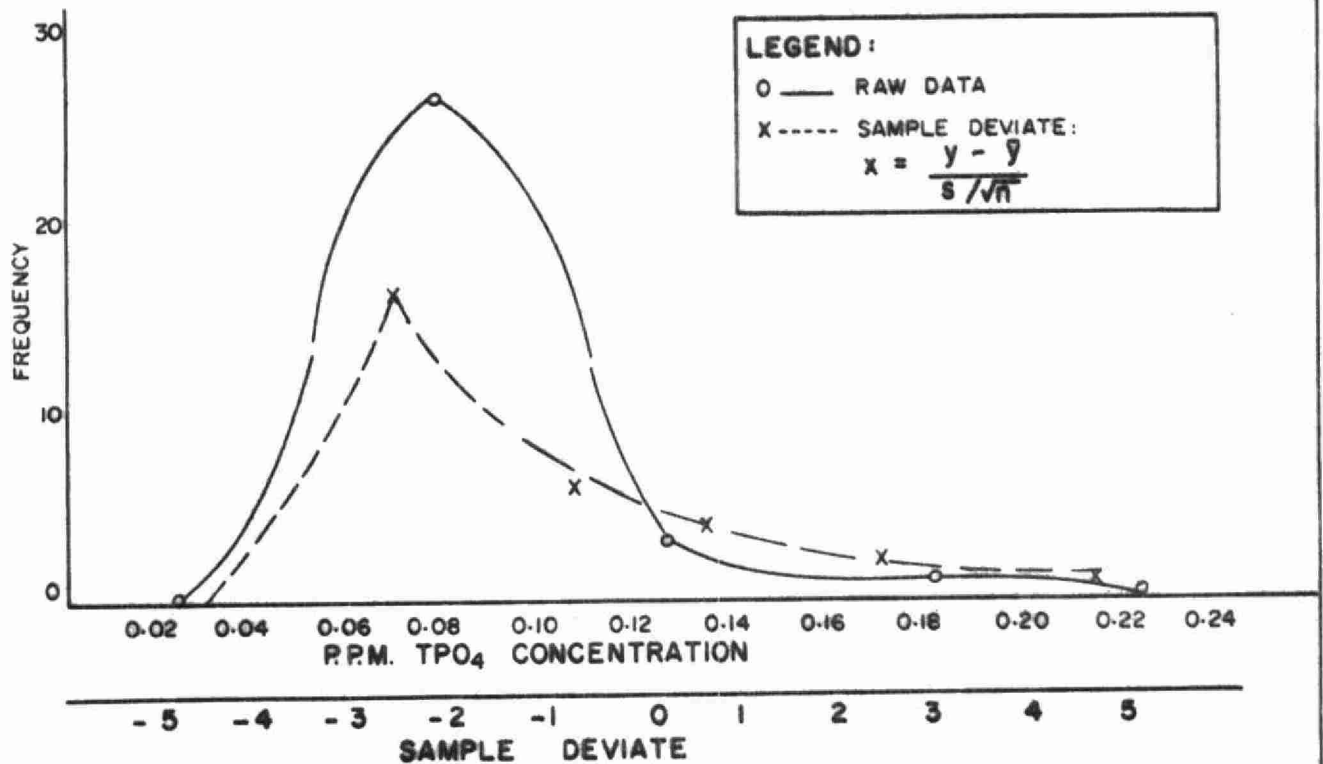


FIGURE 3
ALKALINITY
STATION # 258
1966-67

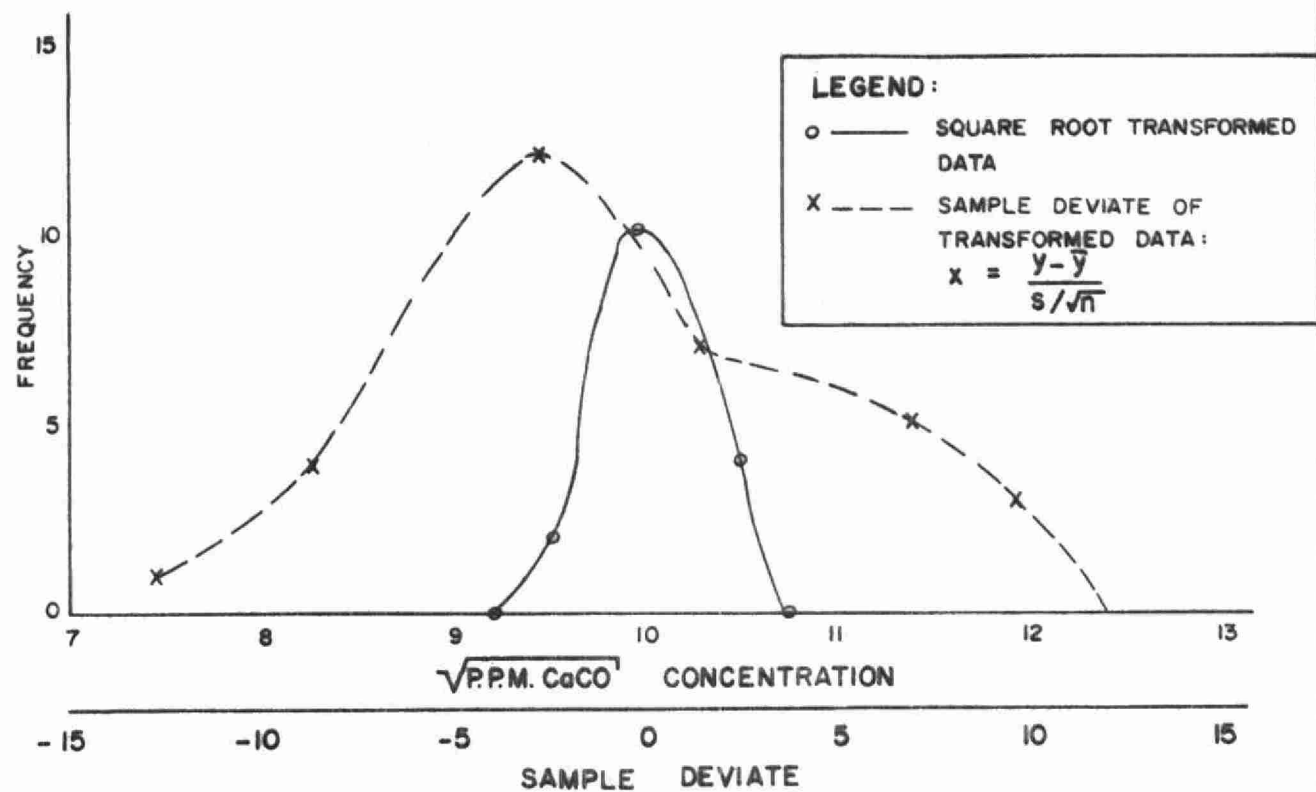
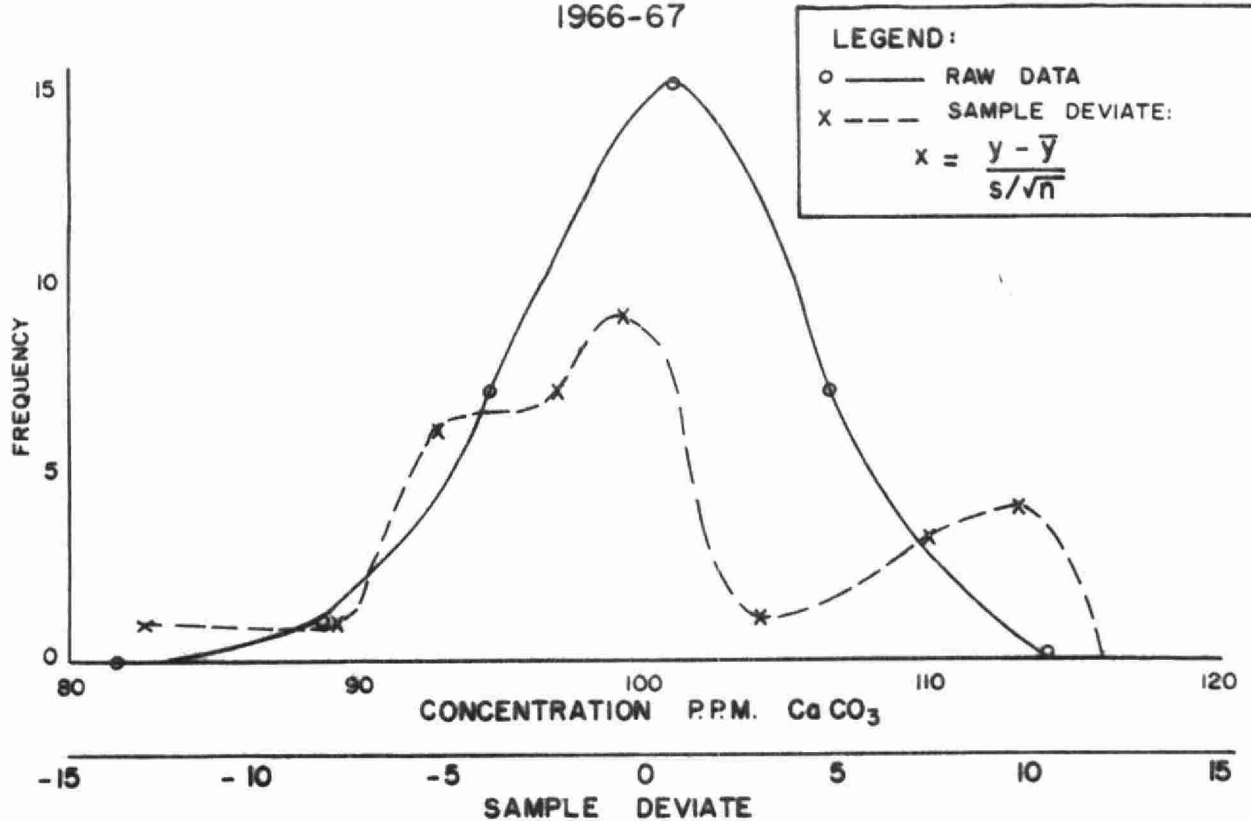


FIGURE 4
DISSOLVED OXYGEN
STATION # 258
1966 - 67

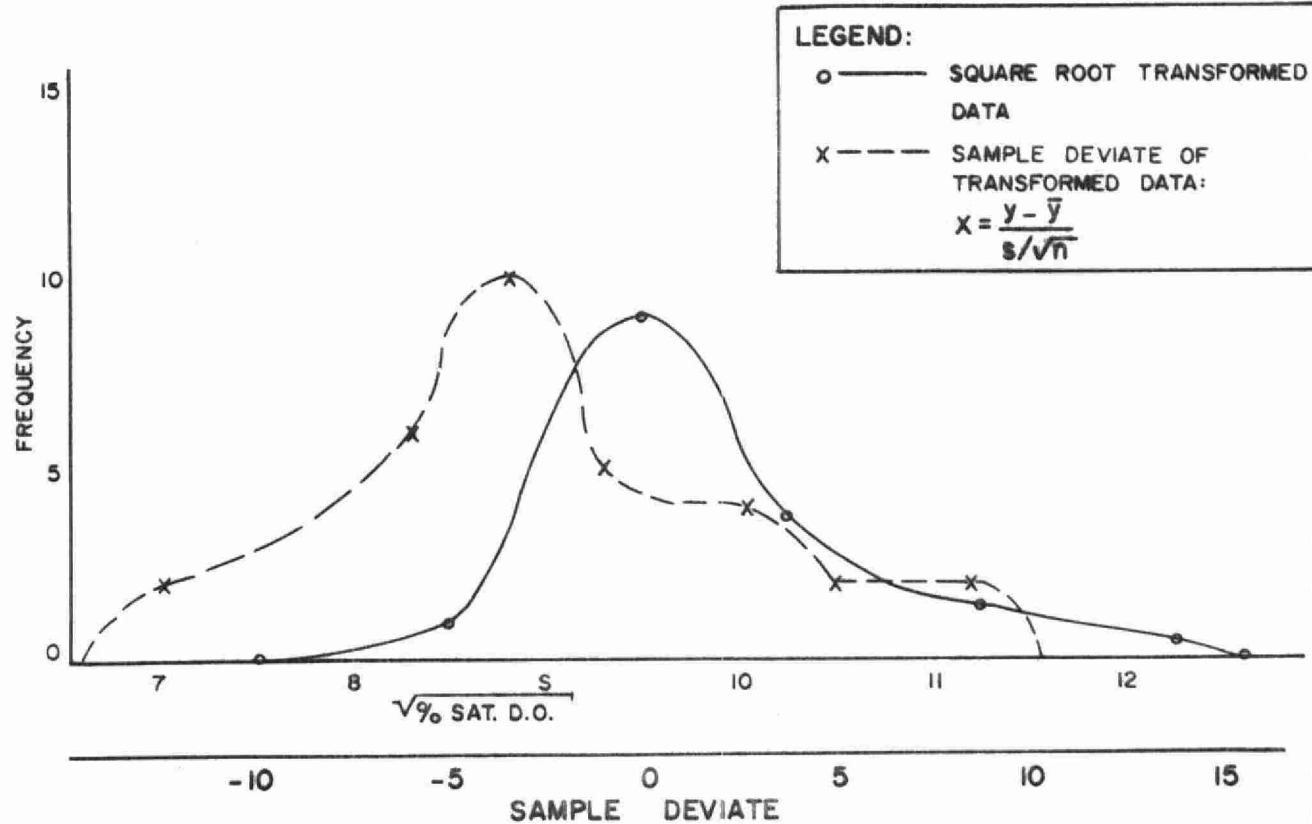
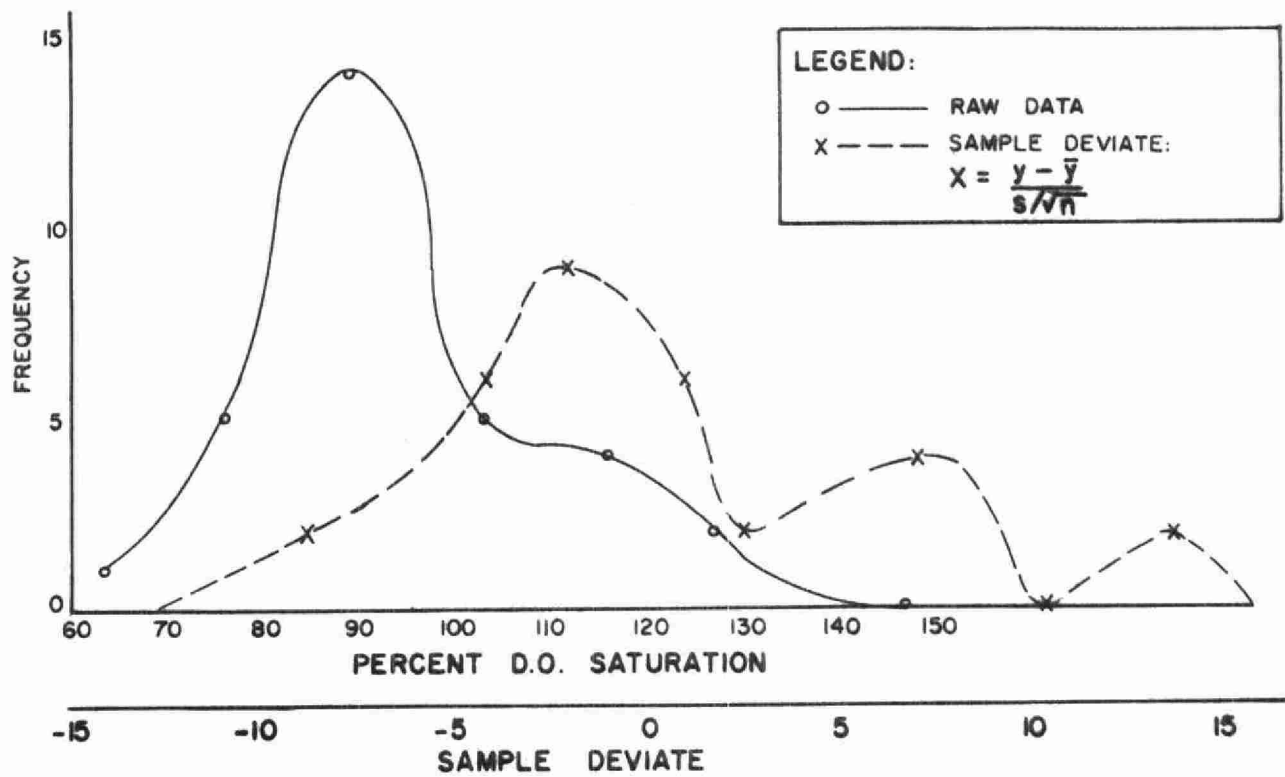
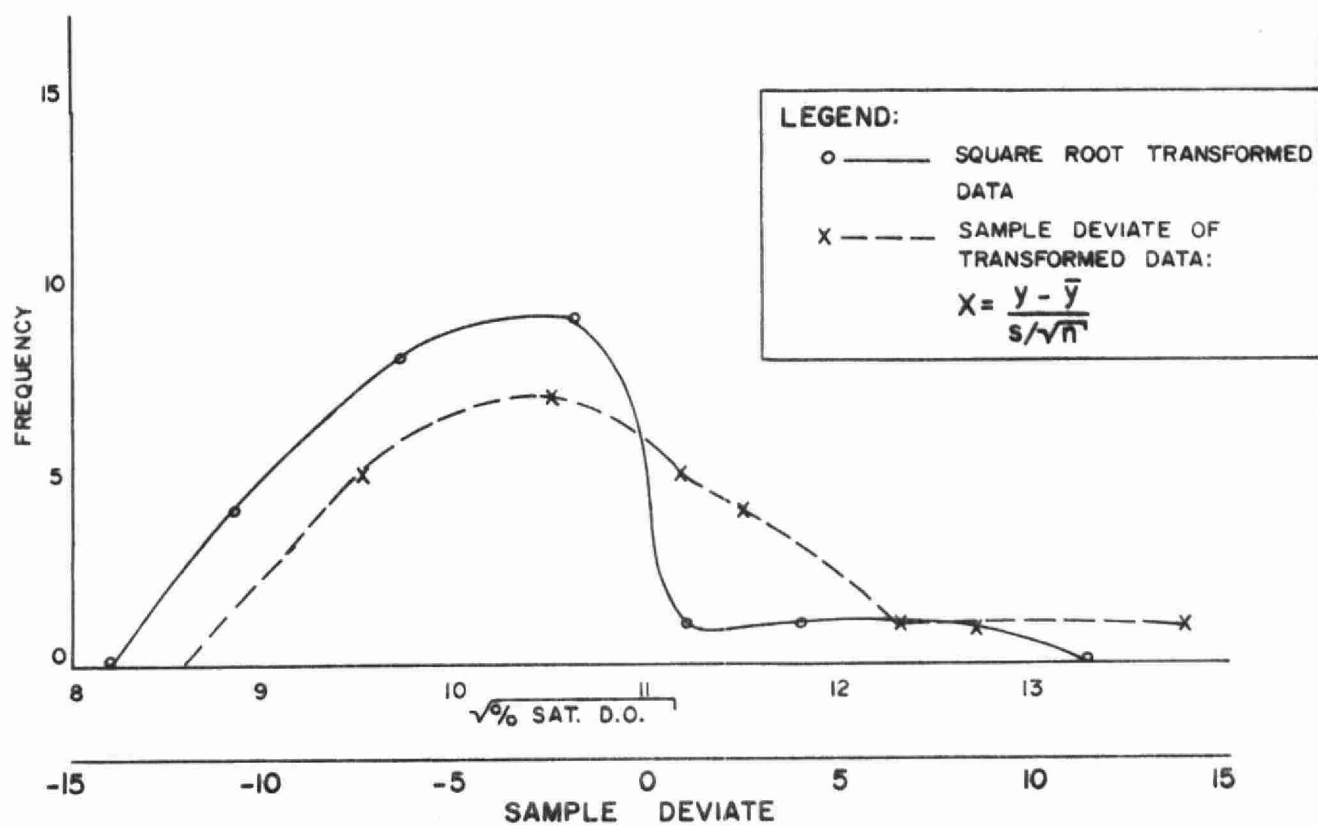
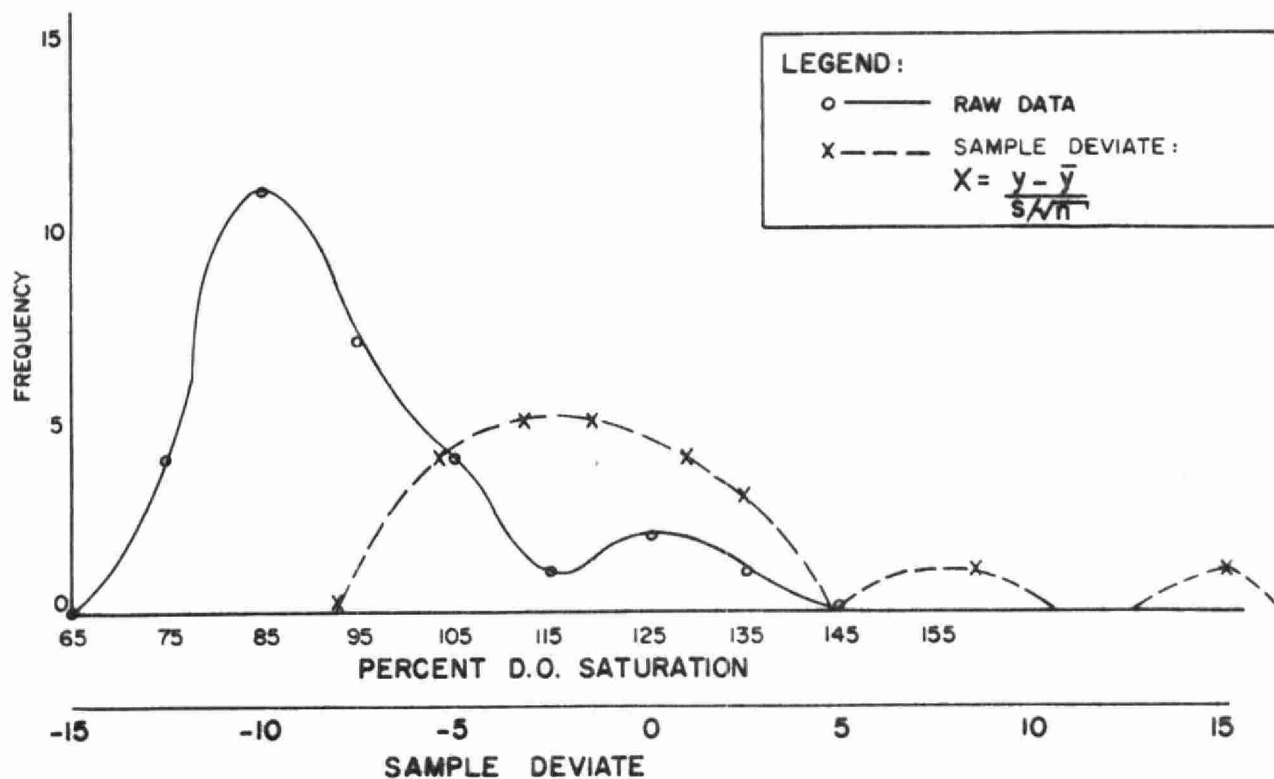


FIGURE 5
DISSOLVED OXYGEN
STATION # 132
1966 - 67



sample. The available number of readings is not sufficient to determine either the distribution or the validity of a transformation. However, there is enough data (approximately 30 readings) to obtain an estimate of the distribution and transformation, particularly if the approximate form of the transformation were known.

Since applicable transformation could not be found in published literature nor through personal contacts (Sayers, 1968), it was necessary to resort to general criteria (U.S. Department of Commerce, 1963, Chapter 20 and Panofsky, 1959, p. 40) for the choice of a transformation. They, however, did not appear to be directly applicable other than the square-root transform which did have some of the required characteristics. The square-root transformation was then applied to the raw data (see figures 2, 3, 4 and 5) and tested parametrically to see if any differences were produced. Although "t" and "F" test values were reduced 4 to 20 per cent, a comparison of the parametric analysis of raw data and transformed data (see Table 3) does not indicate significant changes. In view of the lack of information available on the population distribution or proven transforms by other agencies, it was felt that the searching for a transformation would be a fruitless exercise of guessing with no concrete method of testing the transformation. Consequently, no further transformations were tried. Nevertheless, it should

TABLE 3
PAIR TESTING
PARAMETRIC ANALYSIS F AND T TESTS
RAW AND TRANSFORMED DATA

Stations	No. Readings	Water Quality Parameter	F (calc.)	Remarks	t (calc.)	Remarks
132/140	24/41	DO	1.04	N.S.D.*	0.265	N.S.D.*
		(DO) $\frac{1}{2}$	1.01	N.S.D.	0.251	N.S.D.
	24/31	CaCO ₃	1.17	N.S.D.	-0.145	N.S.D.
		(CaCO ₃) $\frac{1}{2}$	1.11	N.S.D.	-0.131	N.S.D.
132/258	24/32	CaCO ₃	1.27	N.S.D.	-1.13	N.S.D.
		(CaCO ₃) $\frac{1}{2}$	1.43	N.S.D.	-1.6	N.S.D.
258/266	31/31	DO	1.04	N.S.D.	-0.0565	N.S.D.
		(DO) $\frac{1}{2}$	1.00	N.S.D.	-0.0161	N.S.D.
	29/33	T. PO ₄	1.37	N.S.D.	0.586	N.S.D.
		(T. PO ₄) $\frac{1}{2}$	1.25	N.S.D.	0.432	N.S.D.

* N.S.D. - No significant difference

be borne in mind that if such transformations were known, statistical methods could be applied to small samples (only 5-10 values) with confidence. It is suspected that the introduction of recording type water quality meters to lake water quality studies will produce population distributions which will enable the appropriate transformations to be determined.

Non-Parametric Analysis

If a normalizing transformation cannot be found the only alternative is to try non-parametric statistics. Although there is no underlying assumption of normality in non-parametric methods, a larger number of determinations are required to obtain results comparable to the parametric methods. In turn, if parametric methods are applied to non-normal distributions, invalid conclusions can result. Both the Mann-Whitney U and the Kolmogorov-Smirnov two-sample tests (Siegel, 1956 p. 157) are suitable for this case of pair testing. The Mann-Whitney U test is used if one wants to determine whether two samples represent populations which differ in central tendency (mean), while the Kolmogorov-Smirnov two-sample test is used to determine whether two samples are from populations which differ in any respects. The power efficiency (compared to parametric tests) of both these tests is 95 per cent (p. 126 and p. 136, Siegel, 1956).

TABLE 4
PAIR TESTING
NON-PARAMETRIC
MANN-WHITNEY "U" TEST

Stations	Estimated* Frequency Curve Type	No. Readings	Water Quality Parameter	Significance	Probability	Remarks
132/140	NN/NN	24/31	DO	-0.44	0.33	N.S.D.**
	NN/NN	24/31	CaCO ₃	-1.16	0.05	N.S.D.
	N/NN	22/29	S.PO ₄	- .66	0.25	N.S.D.
	N/NN	23/30	T.PO ₄	-1.10	0.14	N.S.D.
132/258	NN/NN	24/31	DO	-1.19	0.028	S.D.
	NN/NN	24/32	CaCO ₃	-1.13	0.13	N.S.D.
	N/NN	22/30	S.PO ₄	- .16	0.44	N.S.D.
	N/NN	23/29	T.PO ₄	-3.25	0.0007	S.D.

* Curve type implies the normal (N) or non-normal (NN) nature of the distribution

** N.S.D. - Not significantly different
- Significantly different

TABLE 4
PAIR TESTING
NON-PARAMETRIC
KOLMOGOROV-SMIRNOV TWO-SAMPLE TEST

Stations	Estimated Frequency Curve Type	No. Readings	Parameter Water Quality	Test K _d Calculated	K _d Two - Tail		Remarks K _d Calculated > K _d S _D
					$\alpha = 0.05$	$\alpha = 0.01$	
132/140	NN/NN	24/31	DO	3	11	13	NSD*
	NN/NN	24/31	CaCO ₃	4	11	13	NSD
	N/NN	22/29	S PO ₄	4	10	12	NSD
	N/NN	23/30	T PO ₄	6	10	12	NSD
132/258	NN/NN	24/31	DO	7	11	13	NSD
	NN/NN	24/32	CaCO ₃	7	11	13	NSD
	N/NN	22/30	S PO ₄	3	10	12	NSD
	N/NN	23/29	T PO ₄	11	10	12	SD

*NSD - not significantly different
SD - significantly different

TABLE 5
PAIR TESTING
PARAMETRIC AND NON-PARAMETRIC
COMPARISON

Stations	Water Quality Parameter	R E S U L T S		
		Parametric	Mann- Whitney "U"	Non-Parametric Kolmogorov- Smirnov 2-Sample
132/140	DO	N S.D.	N.S.D.	N.S.D.
	CaCO ₃	N.S.D.	N.S.D.	N.S.D.
	S.PO ₄	S.D.	N.S.D.	N.S.D.
	T.PO ₄	S.D.	N.S.D.	N.S.D.
132/258	DO	S.D.	S.D.	N.S.D.
	CaCO ₃	N.S.D.	N.S.D.	N.S.D.
	S.PO ₄	S.D.	N S.D.	N.S.D.
	T.PO ₄	S.D.	S.D.	S.D.

TABLE 6

PAIR TESTING

TEST WHETHER THERE IS A SIGNIFICANT DIFFERENCE
BETWEEN TOP AND BOTTOM SAMPLES AT A STATION

PARAMETRIC-T TEST

Station	Water Quality Parameter	No. Readings	t value No information on Variances	t 0.25, n	Remarks
132	DO	9	-2.63	2.31	S.D.*
	Alkalinity	9	0.14	2.31	N.S.D.
	T.PO ₄	9	-1.18	2.31	N.S.D.
258	DO	12	-4.43	2.2	S.D.
	Alkalinity	12	0.21	2.2	N.S.D.
	T.PO ₄	12	0.97	2.2	N.S.D.

* S.D. - Significantly different

N.S.D. - Not significantly different

TABLE 6

PAIR TESTING

TEST WHETHER THERE IS A SIGNIFICANT DIFFERENCE

BETWEEN TOP AND BOTTOM SAMPLES AT A STATION

NON-PARAMETRIC

Station	Water Quality Parameter	No. Readings	Mann-Whitney "U" Test	Kolmogorov-Smirnov Two Sample Test
132	DO	9	N.S.D.*	N.S.D.
	Alkalinity	9	N.S.D.	N.S.D.
	T. PO ₄	9	N.S.D.	N.S.D.
258	DO	12	N.S.D.	S.D.
	Alkalinity	12	N.S.D.	N.S.D.
	T. PO ₄	12	N.S.D.	N.S.D.

* S.D. - Significantly different at $\alpha = 0.05$

N.S.D. - Not significantly different

The results of the non-parametric testing are presented in Table 4 with a comparison of parametric and non-parametric results presented in Table 5. Non-parametric testing shows stations 132/140 to be similar and 132/258 to be different in DO and T PO₄.

Discussion

The difference between testing parametrically and non-parametrically can be seen in Table 5. Parametric testing shows both pairs 132/140 and 132/258 to be significantly different in DO, S PO₄, and T PO₄ whereas non-parametric testing showed 132/140 to be similar in DO, CaCO₃, S PO₄, and T PO₄ and 132/258 to be significantly different in DO, S PO₄ and T PO₄. The different results obtained by the two non-parametric tests is surprising when one considers that the Kolmogorov-Smirnov test is to test for differences in any respects which includes the means tested by the Mann-Whitney U. However, it must be pointed out that S PO₄ under the Kolmogorov-Smirnov is approaching a significant difference at the 0.05 level. The different results produced by the two tests might be due to the nature of the tests. The Mann-Whitney U test is a ranking test which uses a sum of all the ranks in one sample as part of the test statistic whereas the Kolmogorov-Smirnov test statistic is dependent only on the maximum difference in accumulated rank function occurring in a class interval. In other words, one test utilizes all the data available at a station whereas the other is only concerned with one

maximum difference figure. Consequently when one considers all the data available at a station there is a significant difference; however, the maximum difference between the two stations in any class interval is not large enough to be significant. The differences are wider spread over several class intervals. Various numbers of class intervals were tried and were found to affect the results. A similar problem was encountered in the transformation section namely, class interval selection is important. The final rule applied in the selection of class interval has the bulk of the data divided into six to ten classes with isolated extreme values excepted.

Two possible explanations are put forth to account for the conflicting results of the two non-parametric tests. Firstly, the significant diurnal variation must be interacting in some way with the sampling which occurs at various times of the day. Secondly, the thermal stratification effect has not been considered in detail. All samples taken at a station were lumped together in the analysis regardless of the thermal regime. As temperature gradients with depth were not available in sufficient detail, tests were made comparing surface and deep samples (see Table 6) but no significant differences were found. In the multiple station testing section, samples were matched by date and depth in the hope of eliminating the thermal stratification problem.

NB

Conclusions

Testing for station differences should be of the non-parametric type and it is suggested that the Mann-Whitney "U" test be used until more information is available on the conflicting results produced by the Kolmogorov-Smirnov two-sample test, while the parameter which is consistently most sensitive for detecting differences is total phosphorus.

Efforts should be made to determine the population distribution of parameters with recording type meters. Once this is known, the validity of the various forms of testing can be implicitly determined and checked. But even more important, it will provide the necessary information for developing some form of extreme value prediction methods (the maximum concentrations from survey history).

III MULTIPLE STATION TESTING

Extending the pair-testing techniques to multiple station testing is a necessary step for a network of sampling stations analysis. In this case an area containing many sampling stations must be examined to isolate stations which do not add to the water quality information in the area. The analysis of variance provides the vehicle for testing the difference in groups of stations. However, once again, it is necessary to look at both parametric and non-parametric techniques.

Selection of Stations

It was decided to select five stations which would represent areas of low, medium and high values of water quality parameters as delineated by the concentration contouring. The stations selected were as follows:

1. #191
2. #205
3. #211
4. #218
5. #258

Parametric Analysis of Variance

The data from 1966 and 1967 for alkalinity, DO and total PO_4 for the five stations mentioned above were arranged in a two-way

analysis of variance table. In this way, it was possible to individually test for differences between stations and years for each of the selected parameters. The results are presented in Table 8.

One problem of consequence in employing the analysis of variance two-way classification is the requirement for equal number of readings per cell for a fixed effect two-way classification model, (Bowker, 1959, p. 334) unless complex numerical operations are undertaken (Scheffe, 1959, p. 112). This means that if more samples were taken at one station than another it is necessary to select an equal number of samples at all stations. Two commonly used methods are to eliminate samples on the basis of random numbers to the lowest number or to add samples on the basis of some form of extrapolation within the cell to the greatest number. However, it was felt that a superior system existed for equalizing the number of samples in the case of water quality sampling. It was decided to eliminate samples (reduce the number of samples) to conform with the minimum number of samples at a station in the group being considered. This elimination was executed by matching samples at the stations on the basis of depth and date. The data obtained in this way is listed in Table 7.

Analysis of variance (Table 8), once again, indicates that total phosphates proved to be the most sensitive parameter

TABLE 7

MULTIPLE STATION TESTING
PARAMETRIC ANALYSIS OF VARIANCE DATA

	Alkalinity		DO		T.PO ₄	
	1966	1967	1966	1967	1966	1967
191	98	101	118	114	0.50	.40
	103	105	119	111	0.36	.20
	95	96	98	104	.12	.15
	105	103	151	109	.18	.07
	98	108	104	103	.59	.10
	97	108	99	105	.10	.09
	99	101	88	78	.13	.04
	101	95	86	110	.06	.12
205	100	95	114	108	.17	.12
	102	107	116	115	.08	.22
	103	95	105	106	.13	.06
	99	98	112	110	.10	.08
	100	107	98	104	.11	.05
	95	115	104	92	.08	.11
	100	97	89	86	.10	.16
	99	99	87	90	.09	.36
211	102	99	110	100	.32	.14
	100	104	119	119	.14	.08
	100	97	114	106	.26	.23
	100	110	97	125	.13	.10
	100	110	89	104	.37	.09
	103	113	80	105	.07	.11
	100	107	97	89	1.55	.60
	119	101	89	80	.05	.24
218	102	96	121	98	.13	.09
	98	101	140	135	.09	.14
	98	95	123	105	.09	.04
	100	102	99	111	.10	.05
	95	107	75	107	.06	.10
	97	114	76	79	.14	.08
	99	106	111	89	.07	.05
	98	95	86	86	.06	.12

TABLE 7 (Cont'd)

MULTIPLE STATION TESTING

PARAMETRIC ANALYSIS OF VARIANCE-DATA

	Alkalinity		DO		T.PO ₄	
	1966	1967	1966	1967	1966	1967
258	100	97	116	113	.08	.09
	100	96	129	134	.04	.09
	101	99	112	111	.05	.09
	98	100	100	110	.06	.08
	94	97	71	101	.09	.05
	96	106	67	86	.10	.05
	101	108	82	88	.06	.05
	96	94	90	94	.06	.11

TABLE 8

MULTIPLE STATION TESTING
PARAMETRIC ANALYSIS OF VARIANCE

Source	Sum of Squares	Deg. Fred.	Mean Square	Test	F Value from Table	Remarks
Alkalinity:						
Between Years	108.18	1	108.18	4.31	3.98	S.D.
Between Stations	230.53	4	57.60	2.30	2.50	N.S.D.
Interaction	11.63	4	2.91	.12	2.50	N.S.D.
Within combination	1,756.32	70	25.10			
TOTAL:	2,106.66	79	26.60			
Dissolved Oxygen:						
Between Years	19.07	1	19.07	.07	3.98	N.S.D.
Between Stations	303.18	4	75.79	.27	2.50	N.S.D.
Interaction	447.58	4	111.89	.40	2.50	N.S.D.
Within combination	19,826.00	70	283.23			
TOTAL:	20,595.90	79	260.70			
Total Phosphate:						
Between Years	.0437	1	.0437	1.28	3.98	N.S.D.
Between Stations	.4784	4	.1196	3.49	2.50	S.D.
Interaction	.1155	4	.0289	0.84	2.50	N.S.D.
Within combination	2.3980	70	.0343			
TOTAL:	1.3036	79	.0384			

TABLE 9
MULTIPLE STATION TESTING
NON-PARAMETRIC
KRUSKAL-WALLIS ONE-WAY ANALYSIS BY RANKS

Source of Variation	No. Stations or Years	No. Readings per Station or Year	H Statistic	Test H $> \chi^2$	Remarks
<u>Dissolved oxygen</u> between stations	5	16	0.101	$\chi^2 = 0.30$ 0.99;4	N.S.D.
between years 1966-67	2	40	0.07	$\chi^2 = 0.064$ 0.8;1	N.S.D.
<u>Alkalinity</u> between stations	5	16	8.82	$\chi^2 = 7.78$ 0.1;4	N.S.D.
between years 1966-67	2	40	2.18	$\chi^2 = 1.64$ 0.2;1	N.S.D.
<u>Total Phosphate</u> between stations	5	16	21.7	$\chi^2 = 18.46$ 0.001;4	S.D.
between years 1966-67	2	40	0.194	$\chi^2 = 0.15$ 0.7;1	N.S.D.

* Explanation of Test

If the stations are the same, the probability of obtaining an "H" value as high as that tabulated is less than the χ^2 probability. e.g. Alkalinity H=8.82

The probability of obtaining an H this high if the stations are the same $P < 0.10$

For S.D. $P < 0.05$

TABLE 10
 MULTIPLE STATION TESTING
 PARAMETRIC AND NON-PARAMETRIC COMPARISON

Source of Variation	Parametric Analysis of Variance	Non-Parametric Analysis of Variance
ALKALINITY Between Stations Between Years	N.S.D. S.D.	N.S.D.* N.S.D.
DISSOLVED OXYGEN Between Stations Between Years	N.S.D. N.S.D.	N.S.D. N.S.D.
TOTAL PHOSPHATE Between Stations Between Years	S.D. N.S.D.	S.D. N.S.D.

* S.D. - Significantly Different

N.S.D. - Not Significantly Different

for determining station differences. Alkalinity concentrations also showed variations between the two years.

Non-Parametric Analysis of Variance

The choice of analytical techniques is much more limited for multiple station testing than for pair-testing; although the form of the data, a ratio scale, (Siegel, 1956, p. 28) certainly permits a wider choice. It was decided to apply the Kruskal-Wallis one-way analysis of variance by ranks on the same data used for the parametric analysis of variance (Table 7) for comparison of difference between stations and between years. The results of the Kruskal-Wallis analysis are presented in Table 9 and a comparison of the non-parametric and parametric methods for multiple station analysis are presented in Table 10. Total phosphates are once again the most sensitive parameter for indicating differences between stations.

The Kruskal-Wallis method has the ability to handle a variable number of samples at each station and for each year which is not the case for parametric analysis. Furthermore, this method is based upon comparison of averages resulting in an asymptotic efficiency of 95.5 per cent compared to parametric analysis of variance. Consideration was given to the median test but it was felt that the average test would conform better to the concentration

contouring which are founded on averages and would result in a higher efficiency. Unfortunately, no method similar to the Kolmogorov-Smirnov two samples test could be found that would test whether the samples came from populations that differ in any respect at all.

Discussion

The parametric two-way analysis of variance is a powerful well tested method which is capable of testing variations between stations and years or between depths as well as testing for interaction. However, it is necessary to equalize the number of samples in each cell thus introducing another variable of rejecting or fabricating data. While the non-parametric method does not have the limitation of equalizing the number of samples per cell, it is a one-way analysis requiring separate operations for determining station and year differences. Nevertheless, the non-parametric method is nearly as strong as the parametric method.

Furthermore, while proven methods exist for determining specific stations of a group which are different parametrically, non-parametrically such a result can only be achieved by the process of elimination employing successive analysis. This, of course, may not be a serious consideration when a computer is employed.

As in the case of the pair-testing, the problem of diurnal variation and temperature stratification introduces some complications. Matching of dates and depth of samples to account for thermal stratification at each station (see Table 7) did not indicate significant variations. Time of sampling could not be considered to account for the diurnal variation.

It will be noticed in Table 10 that the parametric and non-parametric analyses produce similar differences between stations with the exception of alkalinity. This is surprising when one considers the differences that occurred between the parametric and non-parametric pair-testing. However, it is safer to employ non-parametric methods in multiple station testing because non-parametric methods appear to be better for pair-testing from the point-of-view of consistency until more data is available.

Conclusions

Non-parametric methods similar to the Kruskal-Wallis one-way analysis of variance may be performed on total phosphates to determine if whether station differences in water quality parameters exist as total phosphates were the most sensitive indicator. Further information on diurnal and depth variations is required to streamline analytical techniques. To determine whether differences in other water quality parameters occur more information on the population distributions is required.

Population distributions of various water quality employing recording type meters are required to provide a better understanding of parametric and non-parametric analytical techniques. Once an indication of the form of the population distributions is known, it will be possible to evaluate other analytical techniques.

SUMMARY OF CONCLUSIONS

On the basis of this investigation of pair-testing and multiple station testing the following was determined:

1. Presently the most sensitive water quality parameter for indicating station differences is total phosphate.
2. Presently the best method for determining whether two stations are different is to use the Mann-Whitney U test.
3. Presently the best method for determining whether a group of stations are different is to use the Kruskal-Wallis one-way analysis of variance.

It must be appreciated that these conclusions are temporal. In other words, as more information on the population distributions, thermal stratification, accuracies of testing, and new chemical methods become available, analytical testing methods will change and improve. However, it seems appropriate

to apply the above conclusions to a project area study with some confidence now. Initially, it seems justifiable to ignore diurnal variations and thermal stratification in dealing with the problem of determining whether sampling stations are different. As more data becomes available, they should be checked and incorporated in the analysis.

REFERENCES

- Bowker, A.H. and G.J. Lieberman, Engineering Statistics,
Prentice-Hall Inc., New Jersey, 1959
- Natrella, M.G., Experimental Statistics, National Bureau of
Standards Handbook 91, Washington, D.C. 1963
- Panofsky, H.A., Some Applications of Statistics to Meteorology,
Pennsylvania State University, Pennsylvania, 1965
- Palmer, M.D., Third Water Quality Study Pickering, Ontario Water
Resources Commission, July, 1968.
- Palmer, M.D., Required Density of Water Quality Stations,
Nanticoke, Ontario Water Resources Commission, WQR-1,
1968
- Scheffe, H., The Analysis of Variance, John Wiley and Sons Inc.,
New York, 1959

APPENDIX I COMPUTER PROGRAMS

The following is a list of computer programs which are available for the analysis of data (all programs were tested on textbook examples):

- (1) "FTTEST" is a program used to calculate F values to compare variances of two samples which may have an unequal number of observations. Incorporated into this program are two T-tests to compare means of two distributions. The first one is for the case in which the variances can be assumed to be equal and the second one in which the variances cannot be assumed to be equal.
- (2) "KEN" and "KEN 1" are two programs which contain the basic normalizing transforms outlined in "Experimental Statistics" by M. G. Natrella. Data can be entered into the program and will be transformed according to the transforms in the program.
- (3) Subroutine "FREQ" is a subroutine program written to complement the transformation programs. After the data is transformed, the subroutine can be called directly to give a frequency analysis on the new data.

- (4) "ANOVA" is a one-way classification analysis of variance which will compare factors with the same number or a variable number of readings.
- (5) "ANOVA 1" is a two-way classification analysis of variance for single cell readings.
- (6) "ANOVAT" is a two-way classification analysis of variance for a constant number of multiple cell readings.
- (7) "AVKW" is a one-way non parametric analysis of variance (Kruskal-Wallis).

APPENDIX II
SUMMARY OF ANALYSIS OF PAIRS OF STATIONS
1966 - 67 DATA

Stations	Parameter	No. of Samples	Test	Remarks
258/266	DO	31/31	Parametric	N.S.D.
	T.PO ₄	29/33	Parametric	N.S.D.
191/258	DO	28/31	Parametric	S.D.
	T.PO ₄	27/29	Parametric	S.D.
132/140	DO	24/31	Parametric	N.S.D.
	CaCO ₃	24/31	Parametric	N.S.D.
	S.PO ₄	22/29	Parametric	S.D.
	T.PO ₄	23/30	Parametric	S.D.
132/258	DO	24/31	Parametric	S.D.
	CaCO ₃	24/32	Parametric	N.S.D.
	S.PO ₄	22/30	Parametric	S.D.
	T.PO ₄	23/29	Parametric	S.D.
132/140	DO	24/31	Non-parametric	N.S.D.
	CaCO ₃	24/31	Non-parametric	N.S.D.
	S.PO ₄	22/29	Non-parametric	N.S.D.
	T.PO ₄	23/30	Non-parametric	N.S.D.
132/258	DO	24/31	Non-parametric	S.D.
	CaCO ₃	24/32	Non-parametric	N.S.D.
	S.PO ₄	22/30	Non-parametric	N.S.D.
	T.PO ₄	23/29	Non-parametric	S.D.

APPENDIX IIICHLORIDE (P.P.M.) DATA1966-67

STATIONS				
140	258	132	205	191
1966				
28	27	27	29	27
27	26	27	30	29
26	27	26	27	26
30	27	26	27	26
27	27	26	27	27
28	27	27	27	26
27	27	26	26	27
27	27	27	27	27
26	27	27	26	26
26	26	27	26	27
26	26		27	
26	26		25	
26	27		25	
26	36		26	
26	32			
27	32			
1967				
25	29	26	27	30
24	24	25	27	25
27	24	28	27	27
27	26	26	26	27
27	26	26	27	29
26	27	27	26	27
30	27		28	28
26	26		28	28
28	28		29	28
26	28		29	27
31	27			28
26	27			
28	27			
28	26			

① 2-way analysis of variance and
Choice of data — elimin^g factors.

② " diurnal var^y " observed on
2-2nd runs at 2 different
areas of 2 diff. lakes.
